

Attorney Docket No. 1344.1055

## IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:

Yuichi KAWANISHI, et al.

Application No.: 09/779,495

Filed: February 9, 2001



Group Art Unit: 2857

Examiner:

RECEIVED

JUL 31 2001

TECH CENTER 1600/2900

For: GENETIC MOTIF EXTRACTING AND PROCESSING APPARATUS, GENETIC MOTIF  
EXTRACTING AND PROCESSING METHOD, AND RECORDING MEDIUM  
RECORDED WITH GENETIC MOTIF EXTRACTING AND PROCESSING PROGRAM

**SUBMISSION OF CERTIFIED COPY OF PRIOR FOREIGN  
APPLICATION IN ACCORDANCE  
WITH THE REQUIREMENTS OF 37 C.F.R. §1.55**

Assistant Commissioner for Patents  
Washington, D.C. 20231

Sir:

In accordance with the provisions of 37 C.F.R. §1.55, the applicant(s) submit(s) herewith  
a certified copy of the following foreign application:

Japanese Patent Application No. 10-232901

Filed: August 19, 1998

It is respectfully requested that the applicant(s) be given the benefit of the foreign filing  
date(s) as evidenced by the certified papers attached hereto, in accordance with the  
requirements of 35 U.S.C. §119.

Respectfully submitted,

STAAS &amp; HALSEY LLP

Date: July 25, 2001By: 

James D. Halsey, Jr.  
Registration No. 22,729

700 11th Street, N.W., Ste. 500  
Washington, D.C. 20001  
(202) 434-1500



日 本 国 特 許 庁

PATENT OFFICE  
JAPANESE GOVERNMENT

RECEIVED  
JUL 25 2001  
TECH CENTER 1600/2900

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日  
Date of Application:

1998年 8月19日

RECEIVED

JUL 31 2001

出 願 番 号  
Application Number:

平成10年特許願第232901号

TECH CENTER 1600/2900

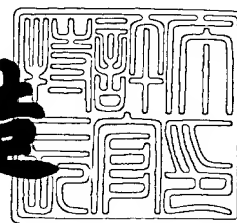
出 願 人  
Applicant (s):

富士通株式会社

2001年 2月23日

特許庁長官  
Commissioner,  
Patent Office

及 川 耕 造



出証番号 出証特2001-3010095

【書類名】 特許願

【整理番号】 9800876

【提出日】 平成10年 8月19日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 17/60  
C12N 15/00

【発明の名称】 遺伝子のモチーフ抽出処理装置及び遺伝子のモチーフ抽出処理方法並びに遺伝子のモチーフ抽出処理プログラムを記録した記録媒体

【請求項の数】 9

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 川西 祐一

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 内藤 公敏

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 宮澤 慎

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 伊藤 辰也

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 二木 敬正

【特許出願人】

【識別番号】 000005223

【氏名又は名称】 富士通株式会社

【代理人】

【識別番号】 100078330

【弁理士】

【氏名又は名称】 笹島 富二雄

【電話番号】 03-3508-9577

【手数料の表示】

【予納台帳番号】 009232

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9719433

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 遺伝子のモチーフ抽出処理装置及び遺伝子のモチーフ抽出処理方法並びに遺伝子のモチーフ抽出処理プログラムを記録した記録媒体

【特許請求の範囲】

【請求項 1】

解明済みの遺伝子配列情報を蓄積する遺伝子配列情報蓄積手段と、

少なくとも 1 つの遺伝子配列情報を入力する遺伝子配列情報入力手段と、

該遺伝子配列情報入力手段により入力された遺伝子配列情報から、遺伝子のモチーフを抽出するモチーフ抽出手段と、

該モチーフ抽出手段により抽出されたモチーフに基づいて、該モチーフを一部に含む遺伝子配列情報を、前記遺伝子配列情報蓄積手段から検索する遺伝子配列情報検索手段と、

該遺伝子配列情報検索手段により検索された遺伝子配列情報を、前記遺伝子配列情報入力手段により入力された遺伝子配列情報に付け加える遺伝子配列情報付加手段と、

を含んで構成されたことを特徴とする遺伝子のモチーフ抽出処理装置。

【請求項 2】

前記遺伝子配列情報入力手段により入力された遺伝子配列情報の中から、前記モチーフ抽出手段によるモチーフの抽出範囲を指定するモチーフ抽出範囲指定手段を備え、

前記モチーフ抽出手段は、前記モチーフ抽出範囲指定手段により指定された抽出範囲内で、遺伝子のモチーフを抽出する構成である請求項 1 記載の遺伝子のモチーフ抽出処理装置。

【請求項 3】

前記遺伝子配列情報を編集する遺伝子配列情報編集手段を備えた構成である請求項 1 又は請求項 2 に記載の遺伝子のモチーフ抽出処理装置。

【請求項 4】

前記モチーフ抽出手段により抽出されたモチーフを編集するモチーフ編集手段を備えた構成である請求項 1 ～請求項 3 のいずれか 1 つに記載の遺伝子のモチーフ

フ抽出処理装置。

【請求項 5】

前記遺伝子配列情報入力手段により入力された複数の遺伝子配列情報を、アラインメント処理するアラインメント手段を備えた構成である請求項 1～請求項 4 のいずれか 1 つに記載の遺伝子のモチーフ抽出処理装置。

【請求項 6】

モチーフを蓄積するモチーフ蓄積手段と、

前記モチーフ抽出手段により抽出されたモチーフを、該モチーフ蓄積手段に登録するモチーフ登録手段と、

を備えた構成である請求項 1～請求項 5 のいずれか 1 つに記載の遺伝子のモチーフ抽出処理装置。

【請求項 7】

前記モチーフ蓄積手段に登録されているモチーフの中から、少なくとも 1 つのモチーフを表示するモチーフ表示手段を備えた構成である請求項 6 記載の遺伝子のモチーフ抽出処理装置。

【請求項 8】

少なくとも 1 つの遺伝子配列情報を入力する遺伝子配列情報入力工程と、

該遺伝子配列情報入力工程により入力された遺伝子配列情報から、遺伝子のモチーフを抽出するモチーフ抽出工程と、

該モチーフ抽出工程により抽出されたモチーフに基づいて、該モチーフを一部に含む遺伝子配列情報を、遺伝子配列情報データベースから検索する遺伝子配列情報検索工程と、

該遺伝子配列情報検索工程により検索された遺伝子配列情報を、前記遺伝子配列情報入力工程により入力された遺伝子配列情報に付け加える遺伝子配列情報付加工程と、

を含んで構成されたことを特徴とする遺伝子のモチーフ抽出処理方法。

【請求項 9】

少なくとも 1 つの遺伝子配列情報を入力する遺伝子配列情報入力機能と、

該遺伝子配列情報入力機能により入力された遺伝子配列情報から、遺伝子のモ

モチーフを抽出するモチーフ抽出機能と、

該モチーフ抽出機能により抽出されたモチーフに基づいて、該モチーフを一部に含む遺伝子配列情報を、遺伝子配列情報データベースから検索する遺伝子配列情報検索機能と、

該遺伝子配列情報検索機能により検索された遺伝子配列情報を、前記遺伝子配列情報入力機能により入力された遺伝子配列情報に付け加える遺伝子配列情報付加機能と、

を実現するための遺伝子のモチーフ抽出処理プログラムを記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、遺伝子のモチーフを自動的に抽出する技術に関し、特に、遺伝子のモチーフ抽出効率を向上させる技術に関する。

【0002】

【従来の技術】

近年の遺伝子工学の進歩に伴い、DNA配列やアミノ酸配列で表現される遺伝子配列の決定技術が急速に進歩している。また、人類を初めとする様々な生物種を対象として、特定生物の遺伝子配列を全て解明しようとするゲノムプロジェクトが世界規模で行われている。このため、解明済みの遺伝子配列情報を有効に活用すべく、遺伝子配列情報データベースが爆発的な勢いで増加している。

【0003】

これらの遺伝子配列の中には、配列情報は明らかになっているが、その機能や構造に関しては未知であるものも多い。かかる遺伝子の機能や構造を配列から推測する有効な方法として、特徴的な規則性を持つモチーフの抽出がある。このため、本願出願人は先の出願において、複数の遺伝子配列情報を比較することで、モチーフを自動的に抽出する技術を提案した（特開平7-274965号公報参照）。

【0004】

【発明が解決しようとする課題】

しかしながら、かかるモチーフ抽出技術では、抽出されたモチーフは、プリンタから出力されるか、或いは、データベース等へファイルとして出力されるだけであって、再利用が充分なされていないのが現状であった。このため、遺伝子の機能や構造の解明を進めるべく、抽出されたモチーフを遺伝子配列情報に加えて、再度モチーフの抽出を行おうとする場合には、抽出されたモチーフ情報を作業者が入力しなければならなかった。従って、従来の技術では、モチーフ抽出効率の向上には自ずから限界があり、さらなるモチーフ抽出効率の向上は困難であった。

## 【0005】

そこで、本発明は以上のような従来の問題点に鑑み、抽出されたモチーフを再利用する機構を付加することにより、遺伝子の機能や構造を推測するためのモチーフ抽出効率を向上させる技術を提供することを目的とする。

## 【0006】

## 【課題を解決するための手段】

このため、請求項1記載の発明は、図1に示すように、解明済みの遺伝子配列情報を蓄積する遺伝子配列情報蓄積手段Aと、少なくとも1つの遺伝子配列情報を入力する遺伝子配列情報入力手段Bと、該遺伝子配列情報入力手段Bにより入力された遺伝子配列情報から、遺伝子のモチーフを抽出するモチーフ抽出手段Cと、該モチーフ抽出手段Cにより抽出されたモチーフに基づいて、該モチーフを一部に含む遺伝子配列情報を、前記遺伝子配列情報蓄積手段Aから検索する遺伝子配列情報検索手段Dと、該遺伝子配列情報検索手段Dにより検索された遺伝子配列情報を、前記遺伝子配列情報入力手段Bにより入力された遺伝子配列情報に付け加える遺伝子配列情報付加手段Eと、を含んで遺伝子のモチーフ抽出処理装置を構成したことを特徴とする。

## 【0007】

かかる構成によれば、遺伝子のモチーフを抽出する場合には、先ず、遺伝子配列情報入力手段を介して、少なくとも1つの遺伝子配列情報を入力する。すると、モチーフ抽出手段により、入力された遺伝子配列情報から遺伝子のモチーフが抽出される。そして、遺伝子配列情報検索手段により、抽出されたモチーフを一



部に含む遺伝子配列情報が遺伝子配列情報蓄積手段から検索される。検索された遺伝子配列情報は、必要に応じて、遺伝子配列情報付加手段により、遺伝子配列情報入力手段により入力された遺伝子配列情報に付け加えられる。その後、モチーフ抽出手段によるモチーフの抽出、遺伝子配列情報検索手段による遺伝子配列情報の検索を繰り返すことによって、遺伝子配列情報の機能や構造を次第に解明していく。即ち、抽出されたモチーフを一部に含む遺伝子配列情報は、遺伝子配列情報付加手段により、入力された遺伝子配列情報に付け加えられるので、抽出されたモチーフの再利用が行われるようになる。

## 【0008】

請求項2記載の発明は、前記遺伝子配列情報入力手段により入力された遺伝子配列情報の中から、前記モチーフ抽出手段によるモチーフの抽出範囲を指定するモチーフ抽出範囲指定手段を備え、前記モチーフ抽出手段は、前記モチーフ抽出範囲指定手段により指定された抽出範囲内で、遺伝子のモチーフを抽出する構成とした。

## 【0009】

かかる構成によれば、モチーフ抽出範囲指定手段によりモチーフの抽出範囲を指定すると、指定された抽出範囲内で、遺伝子のモチーフが抽出される。従って、例えば、入力された種々の生物の遺伝子配列情報の中から、類似する生物の遺伝子配列情報のみを抽出範囲とすることができ、遺伝子配列の機能や構造の解明がより容易になる。

## 【0010】

請求項3記載の発明は、前記遺伝子配列情報を編集する遺伝子配列情報編集手段を備えた構成とした。

かかる構成によれば、遺伝子配列情報に対して、遺伝子配列情報編集手段により各種編集が行われるので、利用者の意思に沿ったモチーフの抽出が行われるようになる。

## 【0011】

請求項4記載の発明は、前記モチーフ抽出手段により抽出されたモチーフを編集するモチーフ編集手段を備えた構成とした。

かかる構成によれば、抽出されたモチーフに対して、モチーフ編集手段により各種編集が行われるので、利用者の意思に沿った遺伝子配列情報の検索が行われるようになる。

## 【0012】

請求項5記載の発明は、前記遺伝子配列情報入力手段により入力された複数の遺伝子配列情報を、アラインメント処理するアラインメント手段を備えた構成とした。

かかる構成によれば、入力された複数の遺伝子配列情報は、アラインメント手段によりアラインメント処理されるので、前記モチーフ抽出手段によるモチーフ抽出が可能になる。また、利用者は、入力された複数の遺伝子配列情報間で類似している領域を視覚的に把握できるため、前記モチーフ抽出範囲指定手段によるモチーフの抽出範囲の指定を容易に行うことができる。

## 【0013】

請求項6記載の発明は、モチーフを蓄積するモチーフ蓄積手段と、前記モチーフ抽出手段により抽出されたモチーフを、該モチーフ蓄積手段に登録するモチーフ登録手段と、を備えた構成とした。

かかる構成によれば、抽出されたモチーフは、モチーフ登録手段によりモチーフ蓄積手段に登録されるため、他の作業者が、抽出されたモチーフを再利用することが可能となる。

## 【0014】

請求項7記載の発明は、前記モチーフ蓄積手段に登録されているモチーフの中から、少なくとも1つのモチーフを表示するモチーフ表示手段を備えた構成とした。

かかる構成によれば、必要に応じて、モチーフ登録手段に登録されているモチーフが表示されるので、種々のモチーフを容易に比較することができ、遺伝子配列の機能や構造の解明がより容易になる。

## 【0015】

請求項8記載の発明は、少なくとも1つの遺伝子配列情報を入力する遺伝子配列情報入力工程と、該遺伝子配列情報入力工程により入力された遺伝子配列情報

から、遺伝子のモチーフを抽出するモチーフ抽出工程と、該モチーフ抽出工程により抽出されたモチーフに基づいて、該モチーフを一部に含む遺伝子配列情報を、遺伝子配列情報データベースから検索する遺伝子配列情報検索工程と、該遺伝子配列情報検索工程により検索された遺伝子配列情報を、前記遺伝子配列情報入力工程により入力された遺伝子配列情報に付け加える遺伝子配列情報付加工程と、を含んで遺伝子のモチーフ抽出処理方法を構成したことを特徴とする。

【0016】

かかる構成によれば、遺伝子のモチーフを抽出する場合には、先ず、遺伝子配列情報入力工程を介して、少なくとも1つの遺伝子配列情報を入力する。すると、モチーフ抽出工程により、入力された遺伝子配列情報から遺伝子のモチーフが抽出される。そして、遺伝子配列情報検索工程により、抽出されたモチーフを一部に含む遺伝子配列情報が遺伝子配列情報データベースから検索される。検索された遺伝子配列情報は、必要に応じて、遺伝子配列情報付加工程により、遺伝子配列情報入力工程により入力された遺伝子配列情報に付け加えられる。その後、モチーフ抽出工程によるモチーフの抽出、遺伝子配列情報検索工程による遺伝子配列情報の検索を繰り返すことによって、遺伝子配列情報の機能や構造を次第に推測していく。即ち、抽出されたモチーフを一部に含む遺伝子配列情報は、遺伝子配列情報付加工程により、入力された遺伝子配列情報に付け加えられるので、抽出されたモチーフの再利用が行われるようになる。

【0017】

請求項9記載の発明は、少なくとも1つの遺伝子配列情報を入力する遺伝子配列情報入力機能と、該遺伝子配列情報入力機能により入力された遺伝子配列情報から、遺伝子のモチーフを抽出するモチーフ抽出機能と、該モチーフ抽出機能により抽出されたモチーフに基づいて、該モチーフを一部に含む遺伝子配列情報を、遺伝子配列情報データベースから検索する遺伝子配列情報検索機能と、該遺伝子配列情報検索機能により検索された遺伝子配列情報を、前記遺伝子配列情報入力機能により入力された遺伝子配列情報に付け加える遺伝子配列情報付加機能と、を実現するための遺伝子のモチーフ抽出処理プログラムを記録媒体に記録した。

## 【0018】

ここで、「記録媒体」とは、電子情報を確実に記録でき、かつ、必要に応じて確実に取り出し可能なものをいい、例えば、磁気テープ、磁気ディスク、磁気ドラム、ICカード、CD-ROM等の可搬媒体が該当する。

かかる構成によれば、遺伝子配列情報入力機能と、モチーフ抽出機能と、遺伝子配列情報検索機能と、遺伝子配列情報付加機能と、を実現するための遺伝子のモチーフ抽出処理プログラムが記録媒体に記録される。従って、かかるプログラムを記録した記録媒体があれば、一般的な電子計算機を用いて、本発明に係る遺伝子のモチーフ抽出処理装置が容易に構築される。

## 【0019】

## 【発明の実施の形態】

以下、添付された図面を参照して本発明を詳述する。

図2は、本発明に係る遺伝子のモチーフ抽出処理装置（以下「モチーフ抽出処理装置」という）の構成を示す。

モチーフ抽出処理装置は、少なくとも、中央処理装置（CPU）とメモリとを備えた電子計算機から構成され、メモリ上にロードされたプログラムに従ってモチーフ抽出に伴う各種処理を実行する。モチーフ抽出処理装置は、データ入力部10（遺伝子配列情報入力手段、工程及び機能）と、データ編集部12と、自動アラインメント部14と、モチーフ抽出部16と、抽出結果処理部18と、モチーフ検索部20と、表示装置22と、印字装置24と、モチーフデータベース26（モチーフ蓄積手段）と、遺伝子配列情報データベース28（遺伝子配列情報蓄積手段）と、を含んで構成される。

## 【0020】

データ入力部10は、例えば、キーボードとマウスとから構成され、モチーフ抽出対象となる少なくとも1つ以上の遺伝子配列データを入力するデータ入力機能を提供する。

データ編集部12は、データ入力部10を介して入力された遺伝子配列データに対して、データ編集機能（遺伝子配列情報編集手段）、抽出範囲指定機能（モチーフ抽出範囲指定手段）、パラメータ設定機能、データ表示機能、外部データ

入力機能及びデータ読込機能を提供する。データ編集機能では、遺伝子配列データに対して各種編集を行うべく、例えば、入力された複数の遺伝子配列データに新たな遺伝子配列データを追加したり、入力された複数の遺伝子配列データから任意の遺伝子配列データを削除したり、任意の遺伝子情報を置き換えたり、手作業によるアラインメント実施機能が提供される。抽出範囲指定機能では、モチーフ抽出範囲を指定すべく、複数の遺伝子配列データからモチーフ抽出対象とする遺伝子配列データを特定したり、アラインメントされた遺伝子情報の範囲を矩形で指定する機能が提供される。パラメータ設定機能では、モチーフの自動抽出を行うための各種パラメータを設定する機能が提供される。データ表示機能では、モチーフ抽出処理装置と利用者とのインターフェースを確立すべく、編集集中の遺伝子配列データを表示装置 22 に表示する機能が提供される。外部データ入力機能では、検索された遺伝子配列の再利用を可能にすべく、モチーフ検索部 20 からのデータを入力する機能が提供される。データ読込機能では、既存のマルチプルアラインメントデータの有効利用を図るべく、ファイルからマルチプルアラインメントデータを読み込む機能が提供される。

#### 【0021】

自動アラインメント部 14 は、データ編集部 12 と協働して、遺伝子配列データを自動アラインメントする機能（アラインメント手段）を提供する。従って、データ編集部 12 に、アラインメント処理済みの遺伝子配列データを入力する必要はなく、例えば、ファイルから遺伝子配列データを複数読み込んだ後、自動アラインメント機能を用いてアラインメント処理を行ってもよい。

#### 【0022】

モチーフ抽出部 16 は、データ編集部 12 の抽出範囲指定機能により指定されたモチーフ抽出範囲に対して、モチーフを自動抽出する自動抽出機能（モチーフ抽出手段、工程及び機能）を提供する。

抽出結果処理部 18 は、モチーフ抽出部 16 により自動抽出されたモチーフ（以下「抽出モチーフ」という）に対して、データベース登録機能（モチーフ登録手段）、データベース参照機能、データ表示機能及びデータ印字機能を提供する。データベース登録機能では、抽出モチーフをモチーフデータベース 26 に登録

する機能が提供される。データベース参照機能では、モチーフデータベース 26 に登録されているモチーフの中から、任意のモチーフを参照する機能が提供される。データ表示機能では、抽出モチーフ或いは参照したモチーフ（以下「参照モチーフ」という）を表示装置 22 に表示する機能が提供される。なお、データベース参照機能及びデータ表示機能により、モチーフ表示手段が構成される。データ印字機能では、抽出モチーフ或いは参照モチーフを印字装置 24 に出力する機能が提供される。また、抽出結果処理部 18 は、データ編集部 12 のデータ編集機能を利用して、抽出モチーフを編集することができるようになっている（モチーフ編集手段）。

#### 【0023】

モチーフ検索部 20 は、遺伝子配列情報データベース 28 から抽出モチーフをその一部に含む遺伝子情報を検索すべく、検索範囲指定機能、検索実行機能（遺伝子配列情報検索手段、工程及び機能）、出力データ指定機能及びデータ表示機能を提供する。検索範囲指定機能では、複数の抽出モチーフの中から、検索を行うモチーフを指定する機能が提供される。検索実行機能では、検索範囲指定機能により指定されたモチーフに基づいて、遺伝子配列情報データベース 28 を検索する機能が提供される。出力データ指定機能では、データ編集部 12 の外部データ入力機能を介して、マルチプルアラインメントデータに追加する検索結果を指定する機能が提供される。なお、データ編集部 12 の外部データ入力機能及びモチーフ検索部 20 の出力データ指定機能により、遺伝子配列情報付加手段、工程及び機能が構成される。データ表示機能では、利用者に対するインターフェースの提供、或いは、検索結果の表示等を行うべく、表示装置 22 にデータを表示する機能が提供される。

#### 【0024】

遺伝子配列情報データベース 28 は、遺伝子配列データベース 28a と、立体構造データベース 28b と、を含んで構成される。遺伝子配列データベース 28a には、解明済みの遺伝子配列情報が登録される。立体構造データベース 28b には、解明済みの遺伝子配列情報であって、その構造を立体的に記述した座標データからなる構造情報が登録される。

## 【0025】

次に、かかる構成からなるモチーフ抽出処理装置の作用について、図3の作業手順図に基づいて説明する。

ステップ1（図では「S1」と略記する。以下同様）では、遺伝子配列データの入力が行われる。

モチーフ抽出処理装置を起動すると、図4に示すような初期画面30が表示される。初期画面30は、データ表示部32と、「ファイル入力」ボタン34と、「自動アラインメント」ボタン36と、「モチーフ抽出」ボタン38と、「データ編集」ボタン40と、「設定」ボタン42と、を含んで構成される。データ表示部32には、入力された遺伝子配列データが表示される。「ファイル入力」ボタン34は、ファイルから遺伝子配列データを読み込むべく、データ読込機能を起動するときにクリックされる。「自動アラインメント」ボタン36は、入力された遺伝子配列データに対してアラインメント処理を行うべく、自動アラインメント部14のアラインメント機能を起動するときにクリックされる。「モチーフ抽出」ボタン38は、入力された遺伝子配列データからモチーフを自動抽出すべく、モチーフ抽出部16の自動抽出機能を起動するときにクリックされる。「データ編集」ボタン40は、入力された遺伝子配列データの編集を行うべく、データ編集機能を起動するときにクリックされる。「設定」ボタン42は、モチーフを抽出する際に必要な各種パラメータの設定を行うべく、パラメータ設定機能を起動するときにクリックされる。

## 【0026】

遺伝子配列データを入力するには、次のような2通りの方法がある。第1の方法は、ファイルから遺伝子配列データを入力する方法である。即ち、初期画面30の「ファイル入力」ボタン34をクリックすると、図示しないファイル指定画面が表示される。そして、ファイル指定画面でファイル名を指定すると、指定したファイルから遺伝子配列データが読み込まれ、図5に示すように、データ表示部32に読み込まれた遺伝子配列データが表示される。また、第2の方法は、データ入力部10のデータ入力機能を使用して、遺伝子配列データを直接入力する方法である。即ち、初期画面30の「データ編集」ボタン40をクリックすると

、図6に示すように、遺伝子配列データの各種編集を行う編集画面50が表示される（編集画面50の詳細は後述する）。そして、データ編集機能の一部を使用して、データ入力部10から遺伝子配列データを直接入力する。

#### 【0027】

ステップ2では、入力された遺伝子配列データの各種編集が行われる。

初期画面30の「データ編集」ボタン40がクリックされると、データ編集部12のデータ編集機能が起動され、図6に示す編集画面50が表示される。編集画面50は、データ表示部52と、「データ追加」ボタン54と、「自動アラインメント」ボタン56と、「モチーフ抽出」ボタン58と、「設定」ボタン60と、を含んで構成される。ここで、データ表示部52、「自動アラインメント」ボタン56、「モチーフ抽出」ボタン58及び「設定」ボタン60は、夫々、初期画面30におけるデータ表示部32、「自動アラインメント」ボタン36、「モチーフ抽出」ボタン38及び「設定」ボタン42と実質的に同一であるので、その説明は省略する。「データ追加」ボタン54は、データ表示部52に表示されている遺伝子配列データに、ファイルから新規に遺伝子配列データを追加すべく、データ読込機能を起動するときにクリックされる。

#### 【0028】

編集画面50上では、例えば、任意の位置にギャップ（図では「-」で示す）を挿入したり、或いは、任意の遺伝子情報を置き換えることができる。また、マルチプルアラインメントデータに新たな遺伝子配列データを付け加える場合には、「データ追加」ボタン54をクリックすると、図示しないファイル指定画面が表示される。そして、ファイル指定画面でファイル名を指定すると、指定したファイルから遺伝子配列データが読み込まれ、データ表示部52に読み込まれた遺伝子配列データが表示される。なお、追加される遺伝子配列データは、マルチプルアラインメントデータであってもよい。

#### 【0029】

ステップ3では、必要に応じて、モチーフ抽出を行うマルチプルアラインメントデータの範囲指定が行われる。

範囲指定を行わない場合には、データ表示部52に表示されたマルチプルアラ



インメントデータ全体がモチーフ抽出対象となるが、実際には、マルチプルアラインメントデータの一部のみをモチーフ抽出対象としたい場合がある。この場合には、抽出範囲指定機能によりモチーフ抽出対象を限定すべく、次のような2通りの方法で範囲指定を行うことができる。第1の方法は、図7に示すように、マウス等のポインティングデバイスを使用して、マルチプルアラインメントデータの中から少なくとも1つのアラインメントデータを指定する方法である。また、第2の方法は、図8に示すように、ポインティングデバイスを使用して、モチーフ抽出対象を矩形領域で指定する方法である。なお、モチーフ抽出対象を変更する場合には、モチーフ抽出対象を指定し直せばよい。

#### 【0030】

ステップ4では、モチーフの自動抽出が行われる。

初期画面30の「モチーフ抽出」ボタン38、或いは、編集画面50の「モチーフ抽出」ボタン58がクリックされると、モチーフ抽出部16の自動抽出機能が起動され、マルチプルアラインメントデータからモチーフが自動的に抽出される。このとき、モチーフ抽出対象が指定されていれば、指定された範囲内からモチーフが抽出される。

#### 【0031】

ステップ5では、モチーフの抽出結果が表示される。

自動的に抽出されたモチーフは、抽出結果処理部18のデータ表示機能により、図9に示すような抽出結果表示画面70に表示される。抽出結果表示画面70は、抽出結果表示部72と、抽出モチーフ表示部74と、「ファイル出力」ボタン76と、「データベース検索」ボタン78と、「アラインメントデータ入力」ボタン80と、を含んで構成される。抽出結果表示部72には、抽出パラメータ72a、解析結果72b等が表示される。抽出モチーフ表示部74には、抽出されたモチーフの一覧が表示される。「ファイル出力」ボタン76は、抽出結果を印字装置24に出力すべくデータ印字機能を起動したり、或いは、抽出結果をモチーフデータベース26に登録すべくデータベース登録機能を起動するときに、クリックされる。「データベース検索」ボタン78は、遺伝子配列情報データベース28に対して検索を行うべく、モチーフ検索部20の検索実行機能を起動す

るときにクリックされる。「アラインメントデータ入力」ボタン80は、現在表示されているアラインメントデータを破棄し、モチーフ抽出を行うアラインメントデータをファイルから読み込む場合にクリックされる。

#### 【0032】

ステップ6では、モチーフ抽出結果の出力が行われる。

即ち、抽出結果表示画面70上に表示されているモチーフ抽出結果を、印字装置24に出力したり、或いは、モチーフデータベース26に登録する処理が行われる。モチーフ抽出結果を出力するには、抽出結果表示画面70の「ファイル出力」ボタン76をクリックすると、出力先を指定する出力先指定画面（図示せず）が表示される。そして、出力先指定画面上で出力先を指定することで、印字装置24へのモチーフ抽出結果の出力、或いは、モチーフデータベース26へのモチーフ抽出結果の登録が行われる。

#### 【0033】

ステップ7では、既存のモチーフ抽出結果との比較が行われる。

即ち、モチーフ抽出結果と、モチーフデータベース26に既に登録されているモチーフ抽出結果と、を比較することにより、モチーフにより推測される遺伝子の機能等の特定作業が行われる。

ステップ8では、遺伝子配列情報データベース28を検索する検索対象の指定が行われる。

#### 【0034】

即ち、図10に示すように、抽出モチーフ表示部74に表示されるモチーフ一覧の中から1つのモチーフをクリックすると、その先頭部に指定されたことを示す記号（図では「\*」で表わされる）74aが付される。検索対象たるモチーフを変更したい場合には、新たなモチーフをクリックすると、そのモチーフが検索対象となる。なお、複数のモチーフを検索対象とするようにしてもよい。

#### 【0035】

ステップ9では、モチーフ検索が行われる。

即ち、検索対象たるモチーフを指定した後で、「データベース検索」ボタン78をクリックすると、指定されたモチーフに基づいて、遺伝子配列情報データベ

ース 28 が検索される。このとき、遺伝子配列情報データベース 28 を構成する遺伝子配列データベース 28 a 或いは立体構造データベース 28 b のどちらを検索するかは、「アラインメントデータ入力」ボタン 80 をクリックすることで任意に選択できる。

#### 【0036】

ステップ 10 では、遺伝子配列情報データベース 28 の検索結果が表示される。

遺伝子配列情報データベース 28 の検索結果は、モチーフ検索部 20 のデータ表示機能により、図 11 に示すような検索結果表示画面 90 に表示される。検索結果表示画面 90 は、モチーフ表示部 92 と、検索結果表示部 94 と、「ファイル出力」ボタン 96 と、「アラインメントへの追加」ボタン 98 と、「エントリ内容表示」ボタン 100 と、「データベース検索」ボタン 102 と、を含んで構成される。モチーフ表示部 92 には、遺伝子配列情報データベース 28 の検索対象たるモチーフが表示される。検索結果表示部 94 には、遺伝子配列情報データベース 28 の検索結果が表示される。「ファイル出力」ボタン 96 は、検索結果を印字装置 24 に出力したり、或いは、モチーフデータベース 26 に登録するときにクリックされる。「アラインメントへの追加」ボタン 98 は、データ編集部 12 の外部データ入力機能を介して、マルチプルアラインメントデータに検索結果を追加するときにクリックされる。「エントリ内容表示」ボタン 100 は、エントリを表示するときにクリックされる。「データベース検索」ボタン 102 は、遺伝子配列情報データベース 28 に対して、絞り込み検索を行うときにクリックされる。

#### 【0037】

ステップ 11 では、再利用する遺伝子配列データが指定される。

即ち、検索結果表示画面 90 において、図 12 に示すように、再利用する遺伝子配列データをクリックすると、指定された遺伝子配列データの先頭部に記号「×」が付される。そして、「アラインメントへの追加」ボタン 98 をクリックすると、データ編集部 12 の外部データ入力機能を介して遺伝子配列データが取り込まれ、図 13 に示すように、マルチプルアラインメントデータの最下段に追加

される。

【0038】

その後、ステップ2～ステップ11を繰り返し、遺伝子の機能や構造を解明する。

かかるモチーフ抽出処理装置によれば、マルチプルアラインメントデータの入力から、モチーフ抽出、抽出モチーフに基づくデータベース検索といった一連の処理を統合化されたシステム上で行うことが可能となる。利用者は、データ編集部12のデータ編集機能を利用することで、入力されたマルチプルアラインメントデータに対して、モチーフ抽出を行うための詳細な設定を行うことが可能となる。また、表示装置22が提供するユーザインターフェースを利用することにより、視覚的に抽出モチーフを確認しながら再度モチーフ抽出を行ったり、或いは、以前に実行した抽出結果との比較を表示装置22上で行うことが可能となる。さらに、モチーフ検索部20の検索実行機能を利用すれば、モチーフ単位でのデータベース検索を容易に行うことができ、また、データ表示機能を利用すれば、検索されたデータのデータベースの登録内容の閲覧が簡単にできる。この結果、遺伝子配列の機能や構造を推測するためのモチーフの抽出を効率良く、かつ、利用者の要求に応じて柔軟的に行うことが可能となる。

【0039】

このような機能を実現するプログラムを、例えば、磁気テープ、磁気ディスク、磁気ドラム、ICカード、CD-ROM等の可搬媒体に記録しておけば、本発明に係るモチーフ抽出処理プログラムを市場に流通させることができる。そして、かかる媒体を取得した者は、一般的な電子計算機システムを利用して、モチーフ抽出処理装置を容易に構築することが可能となる。

【0040】

【発明の効果】

以上説明したように、請求項1又は請求項8に記載の発明によれば、抽出されたモチーフの再利用が行われるようになるので、遺伝子の機能や構造を推測するためのモチーフ抽出効率を向上させることができる。

請求項2記載の発明によれば、モチーフの抽出範囲を指定することができるの

で、例えば、入力された種々の生物の遺伝子配列情報の中から、類似する生物の遺伝子配列情報のみを抽出範囲とすることができ、遺伝子配列情報の機能や構造の解明をより容易に行うことができる。

【0041】

請求項3記載の発明によれば、遺伝子配列情報に対して各種編集が行われるので、利用者の意思に沿ったモチーフの抽出を行うことができる。

請求項4記載の発明によれば、抽出されたモチーフに対して各種編集が行われるので、利用者の意思に沿った遺伝子配列情報の検索を行うことができる。

請求項5記載の発明によれば、入力された複数の遺伝子配列情報はアラインメント処理されるので、必要と思われる遺伝子配列情報を順不同で入力すればよく、遺伝子配列情報の入力作業の効率を向上することができる。

【0042】

請求項6記載の発明によれば、他の作業者が抽出されたモチーフを再利用することが可能となり、遺伝子配列情報の機能や構造を解明する作業の効率をより向上することができる。

請求項7記載の発明によれば、必要に応じて、モチーフを容易に比較することができ、遺伝子配列の機能や構造の解明をより容易に行うことができる。

【0043】

請求項9記載の発明によれば、請求項1又は請求項8に記載の発明の効果に加えて、本発明に係る遺伝子のモチーフ抽出処理プログラムを流通させることができる。従って、かかるプログラムが記録された記録媒体を取得した者は、一般的な電子計算機を用いて、容易に遺伝子のモチーフ抽出処理装置を構築することができる。

【図面の簡単な説明】

【図1】

クレーム1の対応図である。

【図2】

本発明に係る遺伝子のモチーフ抽出処理装置の構成図である。

【図3】

モチーフ抽出を行うための作業手順図である。

【図 4】

初期画面の説明図である。

【図 5】

マルチプルアラインメントデータを入力した初期画面の説明図である。

【図 6】

編集画面の説明図である。

【図 7】

モチーフ抽出範囲を配列で指定した様子の説明図である。

【図 8】

モチーフ抽出範囲を矩形領域で指定した様子の説明図である。

【図 9】

抽出結果表示画面の説明図である。

【図 10】

検索対象を指定した抽出結果表示画面の説明図である。

【図 11】

検索結果表示画面の説明図である。

【図 12】

再利用する配列を指定した検索結果表示画面の説明図である。

【図 13】

再利用する配列がマルチプルアラインメントデータに追加された様子の説明図である。

【符号の説明】

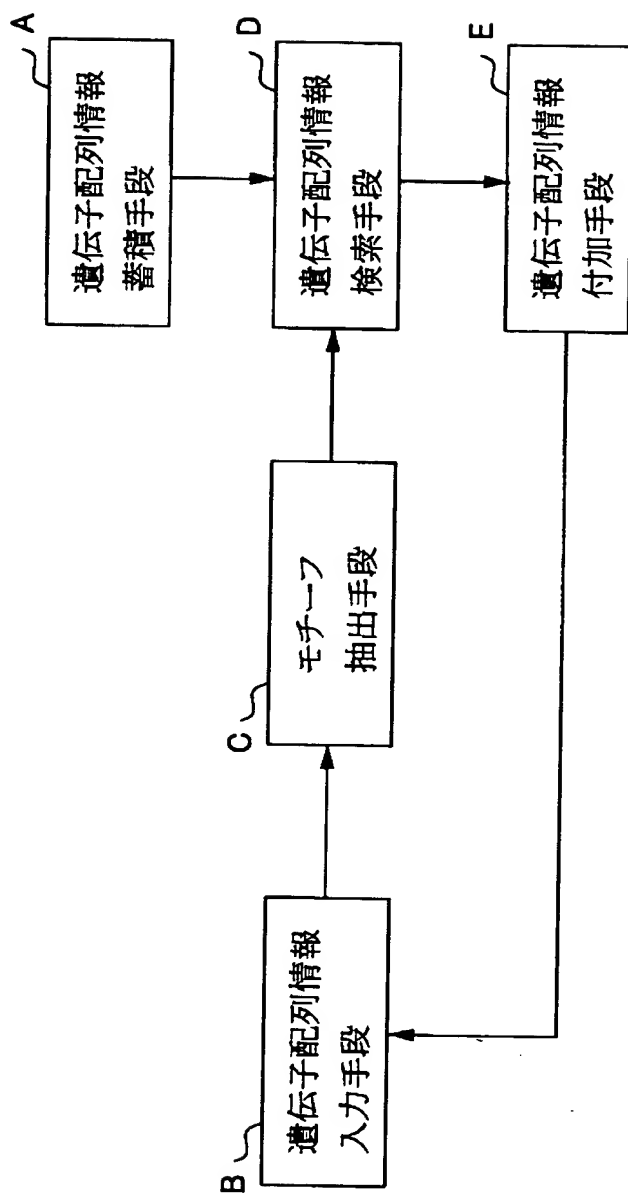
- 10・・・データ入力部
- 12・・・データ編集部
- 14・・・自動アラインメント部
- 16・・・モチーフ抽出部
- 18・・・抽出結果処理部
- 20・・・モチーフ検索部

26・・・モチーフデータベース

28・・・遺伝子配列情報データベース

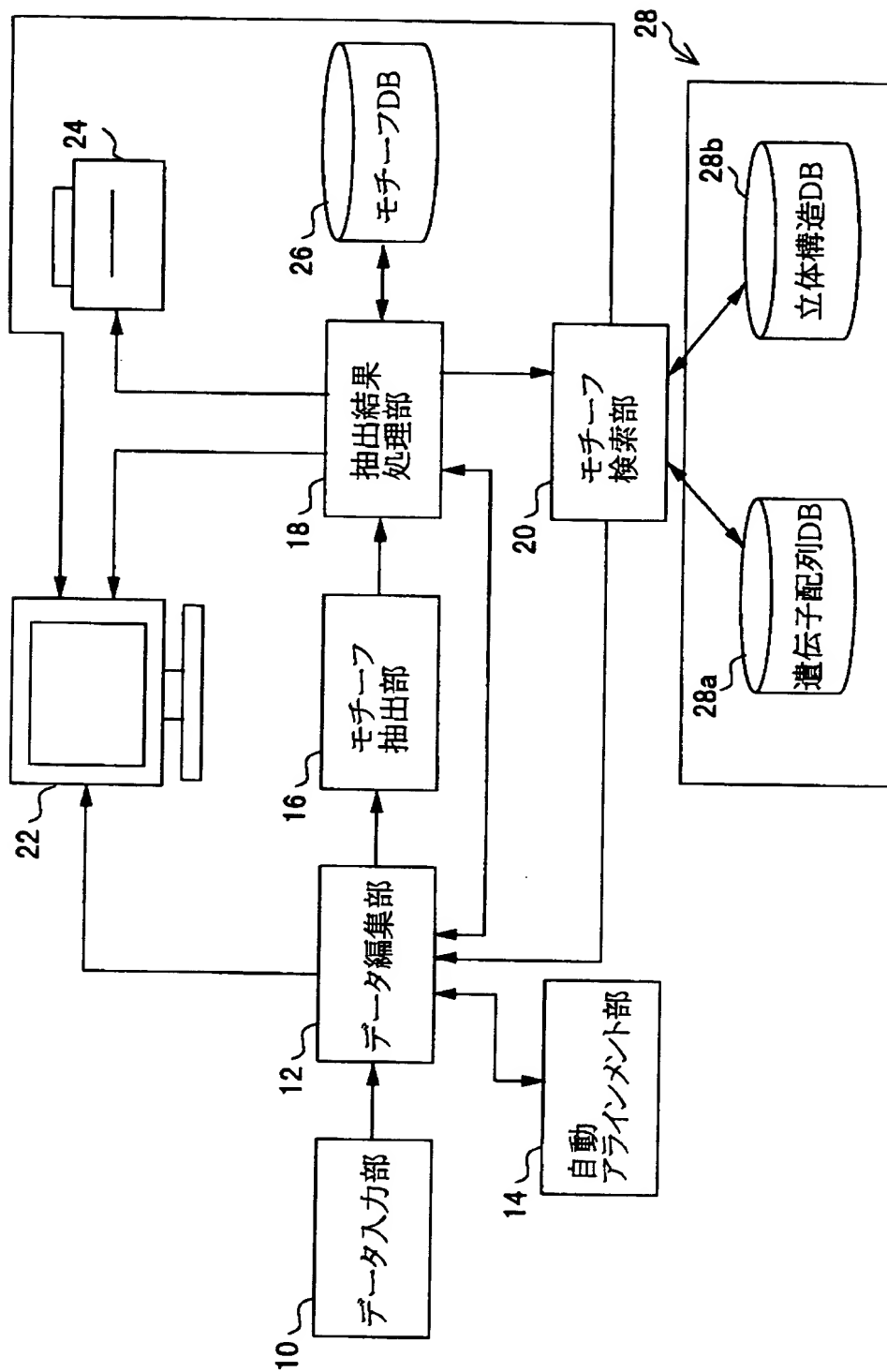
【書類名】 図面

【図 1】

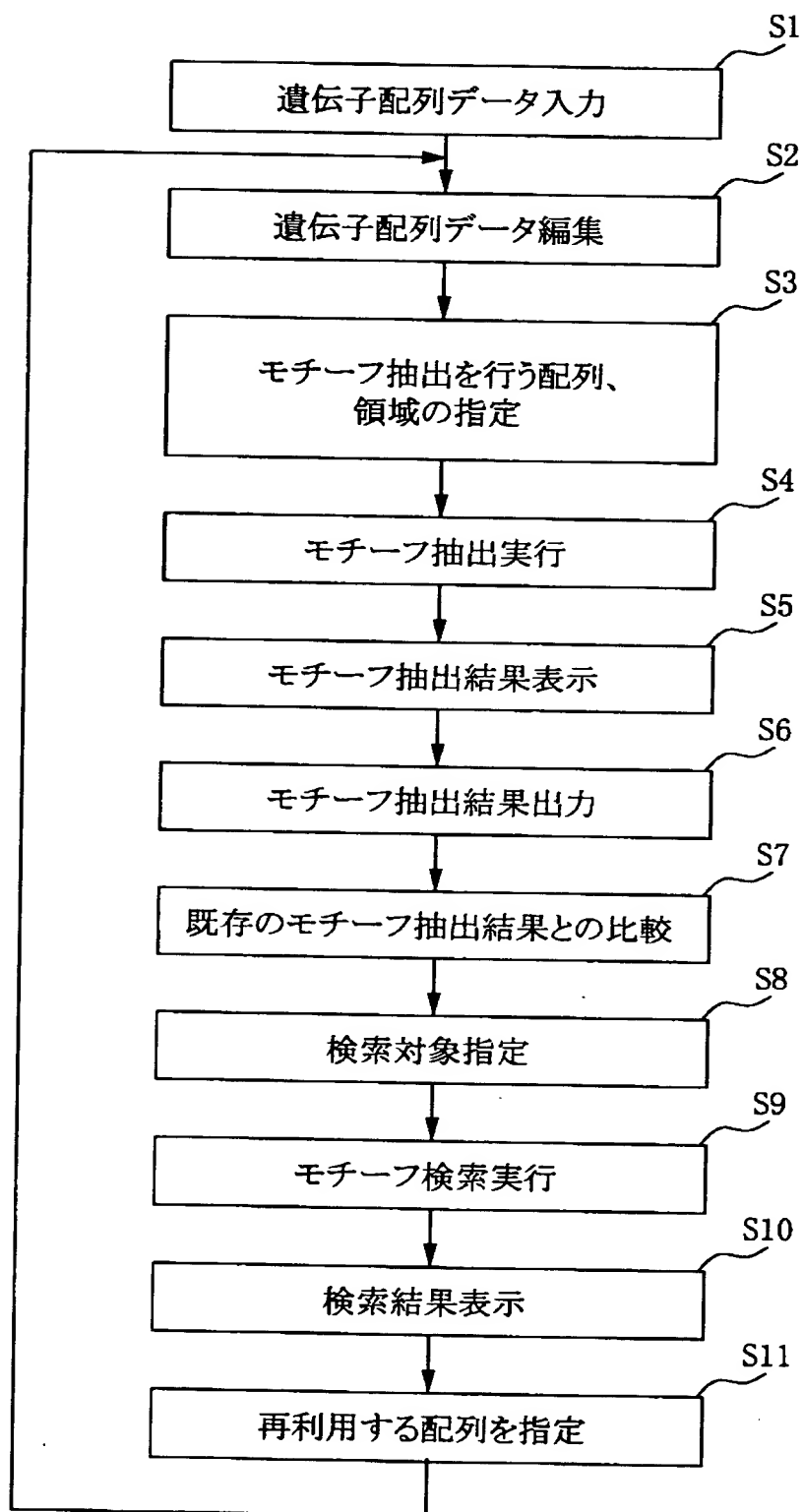




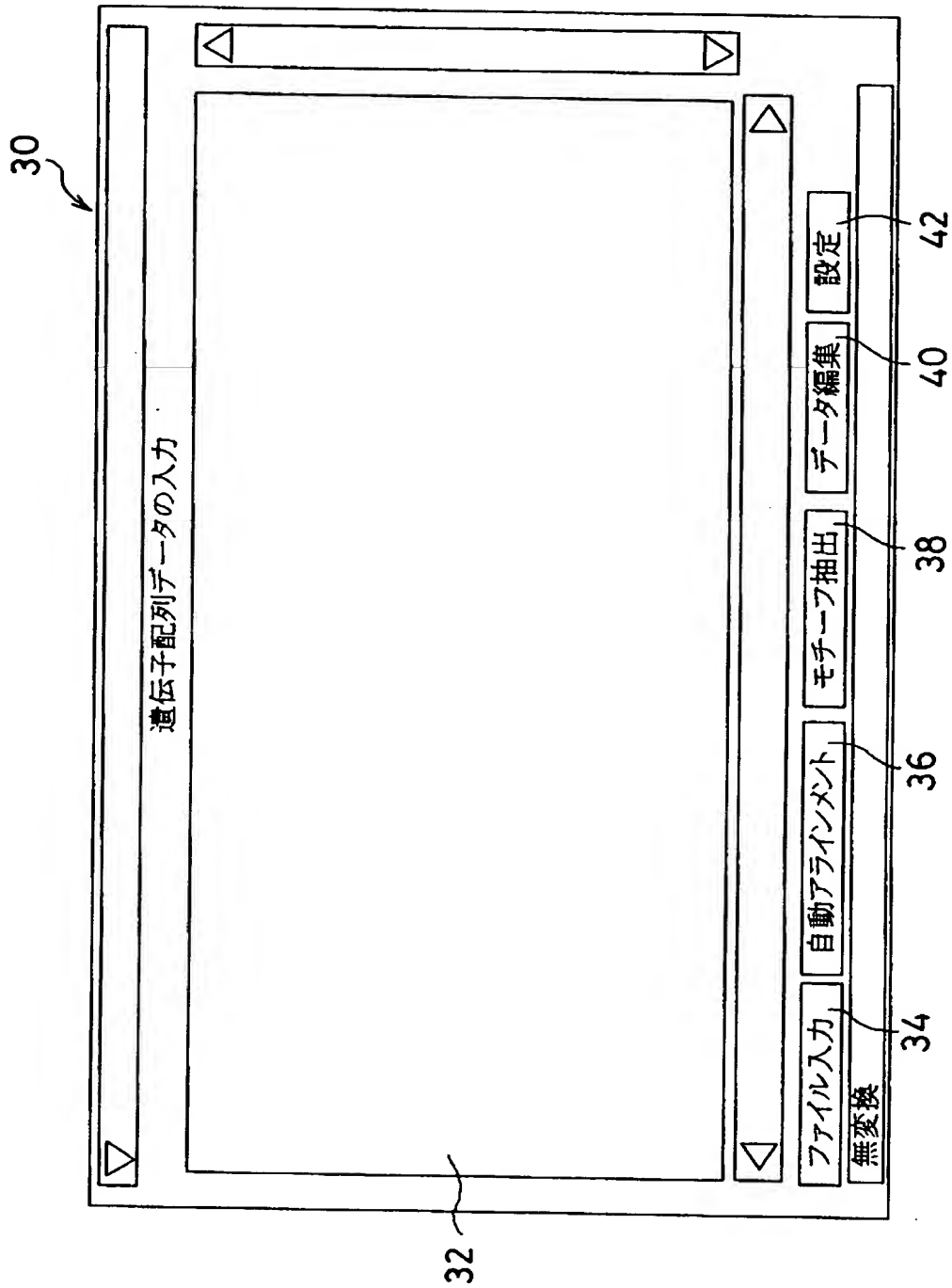
【図 2】



【図 3】



【図4】



【図 5】

30

遺伝子配列データの入力

XLEN2MAB-1  
BFU82487-1  
BRENG1-1  
BRENG2-1  
BRENG2G-1  
BRENG3-1  
MUSEN2AB-1

10 20 30 40 50 60

MEENEQNNREYEPQGE SG EESNRG ILHQAPPG  
MANSADENGNPPASPPPHSPVQPEDSPS  
MEDOR-RQGEEDDSGLSPPLL  
MDENEGSARQVE QR GA D ESNS A IRPLLQ APG  
MDENEGSARQVE QR GA D ESNS A IRPLLQ APG  
MEENDHSNRQVE RQ DS D ESNR A IRPLLQ APG  
MEENDKSPSE-AAEAGRQPEPSSGGSGGSPSDTGRRALMLPEVLQAPGNHGHPR

XLEN2MAB-1  
BFU82487-1  
BRENG1-1  
BRENG2-1  
BRENG2G-1  
BRENG3-1  
MUSEN2AB-1

XLEN2MAB-1  
BFU82487-1  
BRENG1-1

70 80 90 100 110 120

NHDPHRI TNFF IDN LRPEFGRRK  
PDGP RTTNFSIAN LRPEFGARK  
PAHRNTDFF IDN LRPEFG  
NLQLPHRI TNFF IDN LRPEFGRRKEAN I TRYEDN  
NLQLPHRI TNFF IDN LRPEFGRRKEAN I TRYEDN  
NV-LPHRI TNFF IDN LRPEFGRRKEGS RRDEIN  
ITNFF IDN LRPEFGRRKDAGTCCAGGARGGEGAGTTCGGGGGAGGAELDONPTGS

XLEN2MAB-1  
BFU82487-1  
BRENG1-1

XLEN2MAB-1  
BFU82487-1  
BRENG1-1

130 140 150 160 170 180

FTGROT  
RDSKSCSS PVSSPG KPAPG DLSPSLPASPG  
CKR-E RERYTRD SGVRPTALPDSRSORVSSAS

XLEN2MAB-1  
BFU82487-1  
BRENG1-1

ファイル入力

自動アライメント

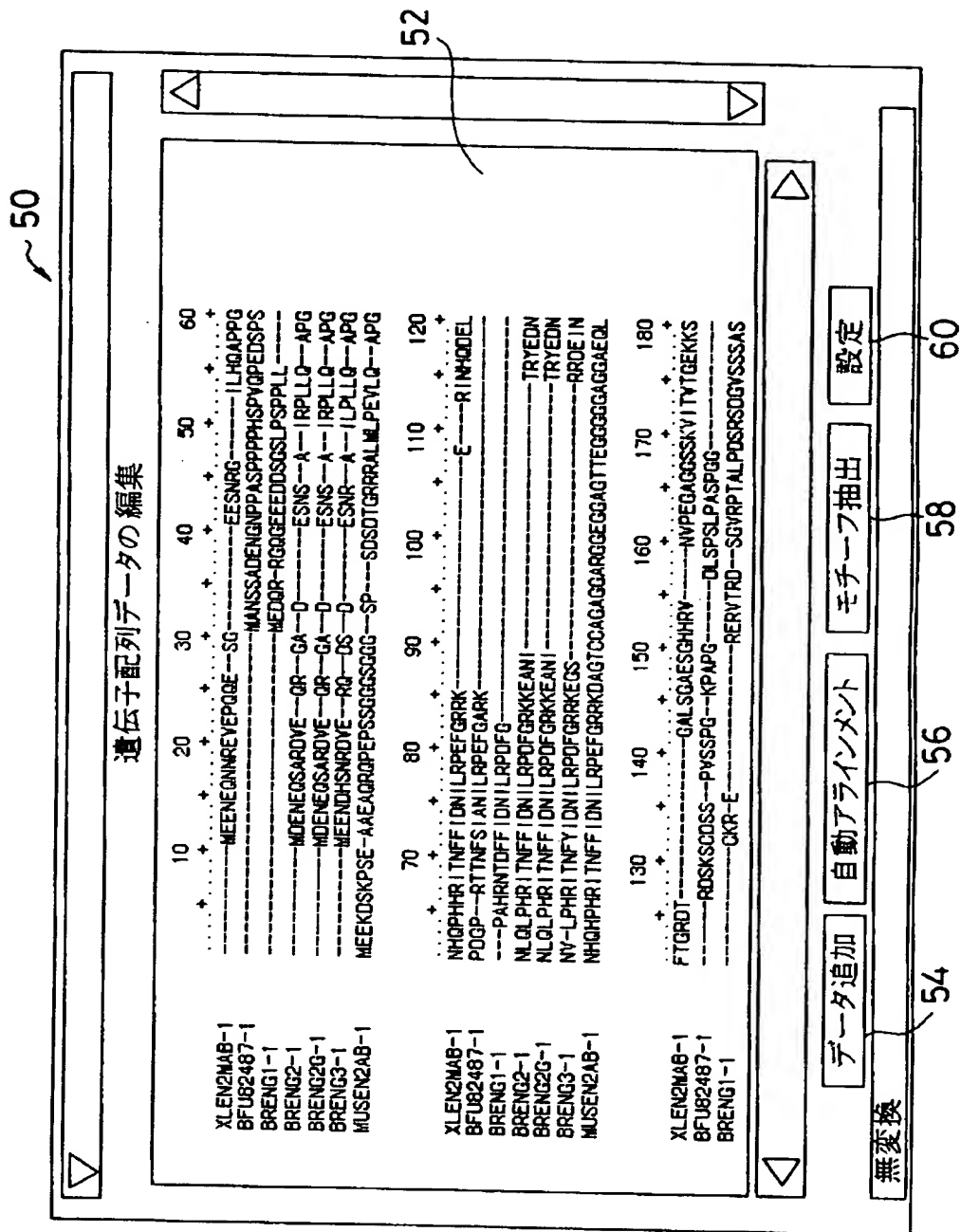
モチーフ抽出

データ編集

設定

無変換

【図 6】



【图7】

55

## 遺伝子配列データの編集集

[illegible]

データ追加  
自動アライメント  
ミチーフ抽出  
設定

無妄

58

【図 8】

50

### 遺伝子配列データの編集

	10	20	30	40	50	60
XLEN2MAB-1	MEENEQNYEYEQE	SG	MANSSAENQNPASP	PPHSPYDPS	PPS	ESNRG
BFU82487-1						ILHAPPG
BRENG1-1						
BRENG2-1						
BRENG2G-1						
BRENG3-1						
MUSEN2AB-1						

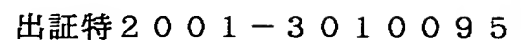
	70	80	90	100	110	120
XLEN2MAB-1	MEENEQNYEYEQE	SG	MANSSAENQNPASP	PPHSPYDPS	PPS	ESNRG
BFU82487-1						ILHAPPG
BRENG1-1						
BRENG2-1						
BRENG2G-1						
BRENG3-1						
MUSEN2AB-1						

	130	140	150	160	170	180
XLEN2MAB-1	FTGRDT					
BFU82487-1						
BRENG1-1						

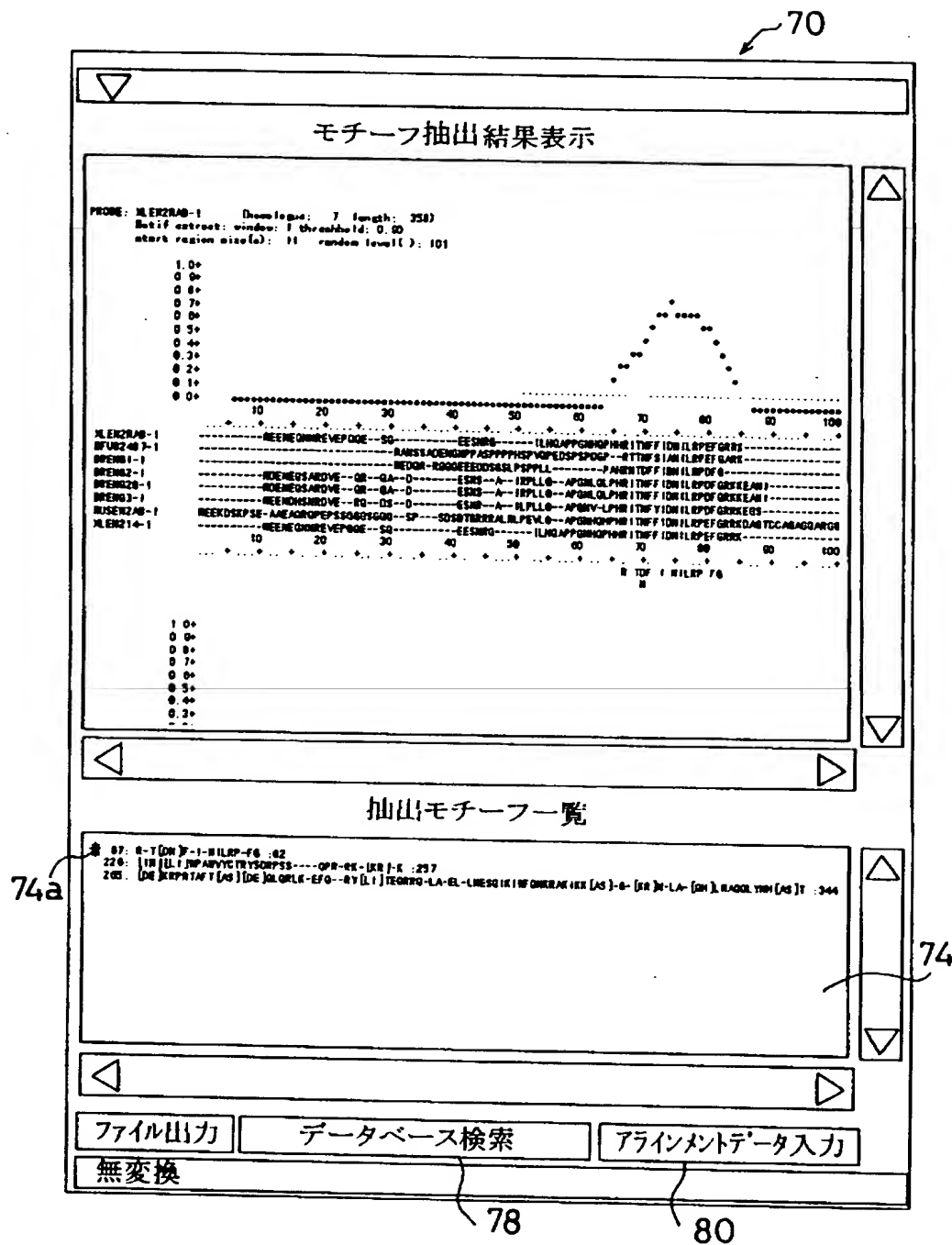
データ追加
自動アラインメント
モチーフ抽出
設定

無変換
58

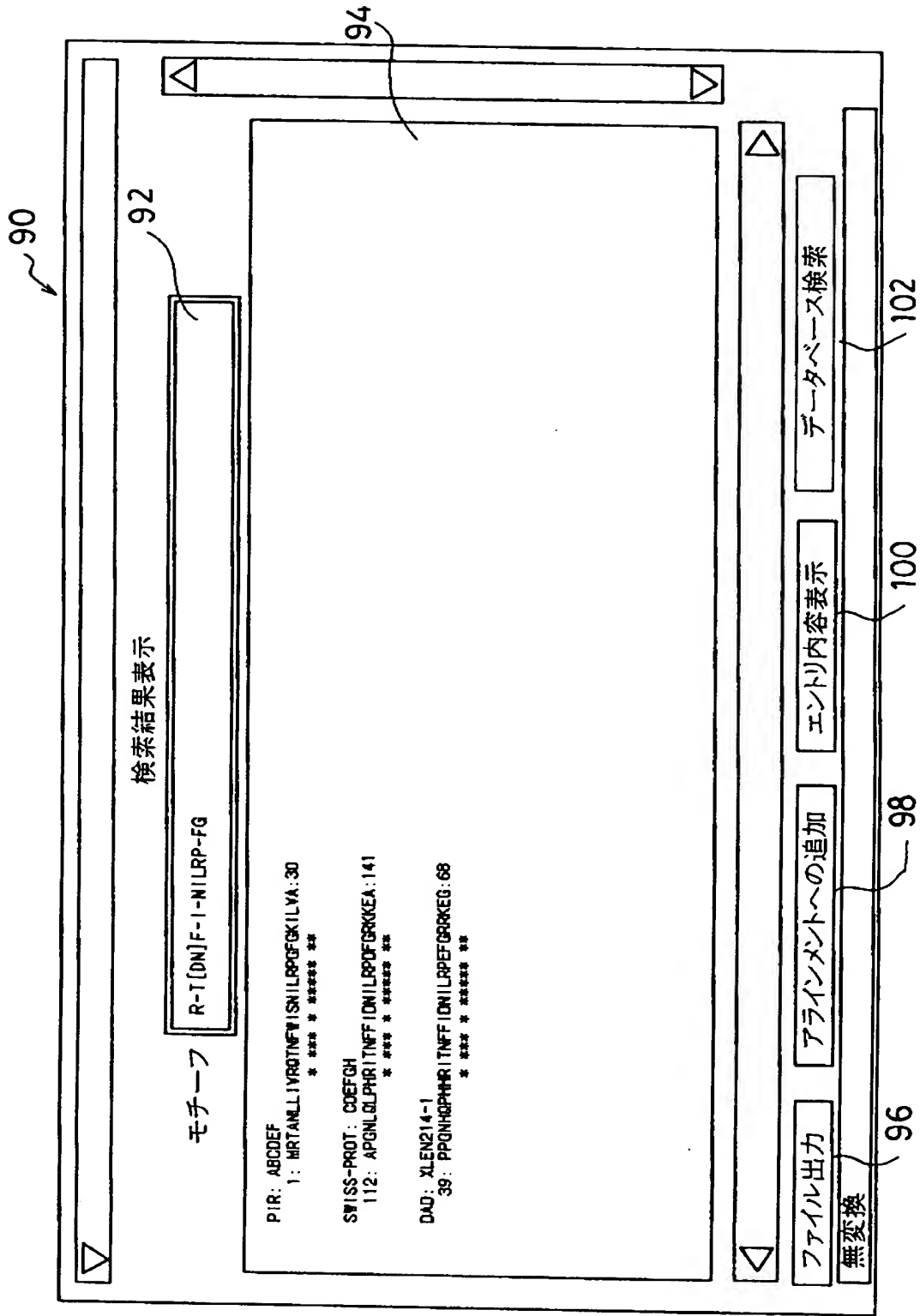




【図 10】



【図 11】



【図 12】

90

▽
△

検索結果表示

モチーフ

R-T[DN]F-I-NILRP-FG

PIR: ABCDEF

1: URTANLLIYRGTNFWISNILRPFGKILVA:30

\* \*\*\* \* \*\*\*\*\* \*\*

SWISS-PROT: CDEFGH

112: APGNLQLPHRIITNFFIDNILRPFGKKEA:141

\* \*\*\* \* \*\*\*\*\* \*\*

X DAO: XLEN214-1

39: PPGNAGPHRIITNFFIDNILRPFGKKEG:68

\* \*\*\* \* \*\*\*\*\* \*\*

▽

△

▽
△

ファイル出力

アラインメントへの追加

エントリ内容表示

データベース検索

無変換

12

出証特 2001-3010095



【書類名】 要約書

【要約】

【課題】 遺伝子のモチーフ抽出効率を向上させる。

【解決手段】 少なくとも1つの遺伝子配列情報を入力するデータ入力部10と、力された遺伝子配列情報に対して各種編集及びモチーフ抽出範囲の指定等を行うデータ編集部12と、入力された遺伝子配列情報或いは指定されたモチーフ抽出範囲から、遺伝子のモチーフを自動的に抽出するモチーフ抽出部16と、抽出されたモチーフの表示及びモチーフデータベース26への登録等を行う抽出結果処理部18と、抽出されたモチーフを一部に含む遺伝子配列情報を遺伝子配列情報データベース28から検索するモチーフ検索部20と、を含んで構成される遺伝子のモチーフ抽出処理装置において、モチーフ検索部20により検索された遺伝子配列情報をデータ編集部12に入力することで、データ入力部10から入力された遺伝子配列情報に付加する機構を付加する。

【選択図】 図2

【書類名】 職権訂正データ  
【訂正書類】 特許願

<認定情報・付加情報>

【特許出願人】

【識別番号】 000005223

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号

【氏名又は名称】 富士通株式会社

【代理人】 申請人

【識別番号】 100078330

【住所又は居所】 東京都港区虎ノ門1丁目19番5号 虎ノ門1丁目  
森ビル 笹島内外特許事務所

【氏名又は名称】 笹島 富二雄

出 願 人 履 歴 情 報

識別番号 [000005223]

1. 変更年月日 1996年 3月26日

[変更理由] 住所変更

住 所 神奈川県川崎市中原区上小田中4丁目1番1号  
氏 名 富士通株式会社